

**INFORMATION TECHNOLOGY  
OPERATIONAL PLAN  
1998-2002**

**FOR  
THE DECENNIAL CENSUS**



**Bureau of the Census  
Economics and Statistics Agency  
Department of Commerce**

**November 7, 1997**

## INTRODUCTION

The purpose of this document is to describe the Information Technology (IT) Operational Plan of the Bureau of the Census to support the Decennial Census which will take place in the year 2000. This document provides a brief description of the Census 2000 and then describes the information technology infrastructure to be used to accomplish the Census. The Plan is organized in the sequence in which the census activities will take place: Data Collection, Data Capture, Data Processing, Data Dissemination. Additionally, there are sections which describe the technology support for Integrated Coverage Measurement, Telephone Questionnaire Assistance, and Administrative Activities. For each of the areas, there is a description of the business activities, the systems required, benefits/justification for the technology solutions, the acquisition strategy, and summary financial data. The Appendix contains, for each area, the major Milestones and the detailed financial information. The Milestones have been extracted from the Master Activity Schedule (MAS), which is used to integrate and manage the multitude of Census 2000 activities. The financial information has been obtained from the Cost Model used for Census 2000. The financial data for each system reflects the requirements as they are currently defined. FY97 is shown as the "Current FY" so that the costs incurred in FY97 can be shown, since that is the first year of major Decennial Census IT investment. The entire Census 2000 is constrained by its overall budget. Therefore, as management decisions are made, the investments in individual systems may vary from the figure shown here. The final section of the Appendix contains a compilation of all the telecommunications costs (which are also reflected in the appropriate system cost data). All of the information is current as of the date of submission. It will be kept current in the Census 2000 MAS and Cost Model, and updated periodically in this plan.

## **1.0 OVERVIEW**

### **1.1 CENSUS 2000**

The major goal of Census 2000 is to perform the best census in our nation's history. In the spring of the year 2000, an estimated 118 million census forms (1.5 billion pages) will be processed in a data capture operation of extraordinary size. This entire operation will be completed during a 99-work-day period. As decreed in the United States Constitution, after this data has been reviewed, processed and tabulated, state level counts used for apportionment must be delivered to the President on or before December 31<sup>st</sup> of the same census year. In addition, within three months of the White House delivery, more detailed state level tabulations and associated maps must be delivered to each state. These state statistics will be used for redistricting.

To build an infrastructure that could support Census 2000 has required many years of planning and foresight. In providing the cornerstone, two core strategies, "Building Partnerships at Every Stage of the Process" and "Using Technology Intelligently" were invoked. Lastly, by linking these strategies to the Census 2000 Acquisition Strategy, it is anticipated that the American public will benefit from the Bureau's partnerships with leading private sector corporations.

Census 2000 will provide the data needed by many agencies and organizations to develop national statistical profiles when defining congressional representation and allocating federal funds. Furthermore, each component of its architecture has met the decision criteria set forth in the Office of Management and Budget (OMB) Memorandum (#M-97-02) dated October 25, 1996 from Franklin D. Raines. Thus, Census 2000 will provide a complete census of the Nation's population and housing, at a lower real cost per housing unit, than the 1990 Census.

### **1.2 CENSUS 2000 PLAN**

The Plan for Census 2000 describes the primary purpose of Census 2000 as follows:

- A complete census of the Nation's population and housing, at a lower real cost per housing unit than the 1990 census.

To accomplish Census 2000, the Plan goes on to describe the following objectives of Census 2000:

- Make every effort to count every household and person--from simpler, user-friendly forms to the better design of field operations.

- Maintain an open process that diverse groups and interests can understand and support.
- Eliminate the differential undercount of racial and ethnic groups.
- Produce a “one-number” census that is right the first time.

The Plan for Census 2000 describes the four strategies for fundamental change, two of which link to the BoC’s strategic use of IT to achieve the objectives of Census 2000.

- **Build Partnership at Every Stage of the Process.** The Census Bureau cannot achieve its goals alone. It needs to award contracts for selected aspects of the process to private sector partners that are “best in class”;
- **Use Technology Intelligently.** Dramatic advances in computing will allow Census 2000 to be simpler, less costly, and more accurate. Through partnership with major contractors, the Census 2000 program will benefit from: the application of the latest technologies; the competence of existing technical staffs; the efficiencies that accrue to firms that continuously operate in the industrial economy; and a more efficient recovery of the short-term technologies used in the system’s infrastructure.

### 1.3 CENSUS 2000 OPERATIONAL AREAS

- **Data Collection:** Supports all field data collection activities for Census 2000. The overall objective of data collection is to obtain a questionnaire from every housing unit in Census 2000. Once the data has been collected, the forms will be sent to the Data Capture Centers (DCCs).
- **Data Capture:** The Census 2000 data capture methodology will utilize imaging technology that will accommodate the use of respondent-friendly questionnaires. Once the data has been captured, an ASCII file will be electronically transmitted to data processing.
- **Data Processing:** The objective is to develop an effective and efficient automated data processing system for controlling, managing, and processing Census 2000 data. The Census 2000 Data Processing System will be a complex network of operational controls and processing routines intended to store and service the decennial control and data requirements. Once the data has been processed, Census 2000 data will be electronically transmitted to the Data Access and Dissemination System (DADS).

- **Data Dissemination:** Census 2000 data will be tabulated and disseminated using the Data Access and Dissemination System (DADS). DADS will provide an interactive electronic system that will allow data users to access prepackaged data products, data documentation, and on-line help, as well as build custom data products on-line and off-line.
- **Integrated Coverage Measurement:** The objective of the Integrated Coverage Measurement (ICM) survey is to produce a “one number” census estimate of the U.S. population in Census 2000 that will improve accuracy, reduce costs, and eliminate confusion and controversy caused by having more than one number measuring the same population.
- **Telephone Questionnaire Assistance:** Telephone Questionnaire Assistance (TQA) will provide respondents with convenient access to obtain assistance in the completion of their Census 2000 questionnaires or to respond directly to the census.
- **Administrative Systems:** The Census Bureau has instituted several management systems to facilitate a more effective and efficient planning process for Census 2000 as well as its actual implementation.

#### 1.4 1998 DECENNIAL DRESS REHEARSAL

The purpose of the 1998 Census 2000 Dress Rehearsal is to test all the various operations, systems, procedures, and questions that are planned for Census 2000 under as near census-like conditions as possible. The Dress Rehearsal will provide for operational testing of Regional Census Center, Local Census Office Data Capture Center, and Headquarters processing procedures and systems planned for use in Census 2000. We will employ the full array of methods, techniques, materials, workflows, equipment applications, and promotion and outreach programs intended for use in Census 2000.

The results of the Dress Rehearsal will be assessed in the Census 2000 Dress Rehearsal Evaluation Program. That Program will define the evaluation studies and quality assurance reviews to be applied to the Dress Rehearsal. The evaluations required for the information technology systems used in the Dress Rehearsal will be documented as part of that program, as will the results. The specific tests to be accomplished and the decisions to be based on them will be part of the evaluation study. The results will be used for future decision making and for any system changes which may be required for the Census 2000.

## 2.0 DATA COLLECTION

### 2.1 Description

The data collection system supports all field data collection activities for Census 2000. The overall objective of data collection is to obtain a questionnaire from every housing unit in Census 2000. However, to achieve this objective there are a number of activities which must be completed. These activities have been broken into the following two data collection phases:

- **Early Census Field Operations.** Starting two years prior to Census Day operations, certain address compilation operations must occur all over the country to create and update address lists that will be used to deliver census questionnaires for Census 2000. These early census field operations require a large national scale infrastructure in order to complete the mission.
- **Census Field Operations.** Census Field Operations includes field data collection, where an enumerator visits the household. These operations begin and continue through 2000. The census field operations, like the early census operation, require a large national scale infrastructure in order to complete the mission.

In order to set up an effective and cost-efficient infrastructure that facilitates the complete and accurate enumeration in Census 2000, the Census Bureau will develop an extensive temporary office infrastructure across the country. The major field infrastructure components are as follows:

**Regional Census Centers (RCCs)** consist of 12 stateside Centers and an Area Office in Puerto Rico, these offices will begin opening in November of 1997 and close in early 2001. RCCs are large, temporary offices designed to provide regional coordination for both early census field operations in the Census Field Offices (CFOs), and census field operations in the Local Census Offices (LCOs). The RCCs will house high-end Unix servers, responsible for storing and processing data for most of the data collection subsystems. The RCC's will provide the following coordination and support for both early census and census field operations:

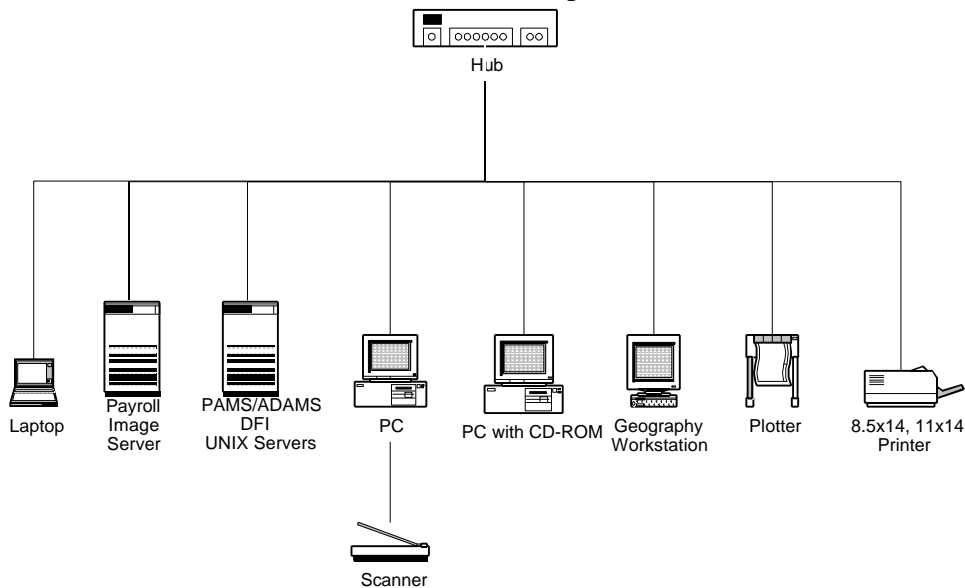
- Provide overall management of Census Field Operations;
- Manage all address listing, and address list enhancement for city style address areas through a network of Census Field Offices (CFOs);
- Produce maps (this operation is covered in detail in the Geography IT Operational Plan);
- Conduct geographic activities such as geocoding, and TIGER data base updates (these operations are covered in detail in the Geography IT Operational Plan);
- Conduct recruiting;

- Manage payroll and personnel administrative system.

**Census Field Office (CFOs)** consist of 425 to 450 Offices that will begin opening in November of 1997 with the last office closing in December of 1998. CFOs are very small, temporary offices designed to support the major early census field operations. There will be no automation in the CFOs. Instead, all work elements scheduled for the CFOs will be automated by the RCCs. Assignments will be generated within the RCCs and distributed to the CFOs via USPS and over night mailings.

**Local Census Offices (LCOs)** consist of 475 Offices (466 stateside, 9 in Puerto Rico) that will begin opening in the Fall of 1999 and close in the Fall of 2000. LCOs are small, temporary offices designed to support the major census field operations.

**RCC Architecture** The following architecture will fully support both early census and census field operations through 2001. Within each RCC, three high-end Unix servers will be responsible for storing and processing data for most of the Decennial Field Interface (DFI) subsystems. Users will access these servers over an 802.3 10BaseT Local Area Network (LAN) (see below) using PC-based (Windows 95) desktop computers (40 PCs). Each RCC will also contain digitizing workstations (6 PCs), imaging workstations (3 PCs), electrostatic plotters (10) and networked printers. The RCCs will be connected to the BoC enterprise network over T1 or T3 Frame Relay links.



**RCC ARCHITECTURE**

## 2.1.1 Field Operations

**Address Listing Operation** One of the major functions to be performed during Early Census Field Operations is Address Listing. This operation, which is conducted primarily in rural areas, will construct a comprehensive address list of approximately twenty-two million housing units and special places by systematically canvassing streets, trails, etc., located within each block in an assignment area. For each housing unit, the enumerator will attempt to obtain the physical location address or description, mailing address, and the household name from an occupant of the housing unit or other knowledgeable person. In addition, the enumerator will assign each housing unit to the census collection block in which it is located, and update the block maps with the housing unit location. While this operation will be conducted out of the CFOs, actual assignments will be generated within the RCCs and distributed to the CFOs.

**Census Field Office (CFOs)** The CFOs will provide the following support during early census field operations:

- Perform address listing or address list enhancement for city style address areas;
- Conduct local recruiting;
- Perform clerical review of completed field work.

**CFO Architecture** There will be no automation in the CFOs. Instead, all work elements scheduled for the CFOs will be automated by the RCCs.

## **2.1.2 Census Field Operations**

To ensure that we obtain a completed questionnaire from every housing unit, the Census Bureau must first make sure that a questionnaire is delivered to every housing unit. Field data collection (where an enumerator visits the household) will accomplish this by using one of the following two basic data collection methods:

1. **Update/leave.** In areas where there is no USPS delivery to city style addresses, enumerators will update the address list through canvassing and drop off addressed census questionnaires to each housing unit for the householder to complete and mail back;
2. **List/enumerate.** In very remote or sparsely-populated areas, enumerators will visit each housing unit and pick up or complete unaddressed short-form questionnaires that the USPS previously delivered to each unit.

In addition, the Census Bureau will implement a comprehensive set of procedures to enumerate people who do not live in traditional housing units. These include

people who live in group quarters situations (for example, nursing homes, group homes, and colleges), people without housing, and people who live at migrant and seasonal farm worker camps.

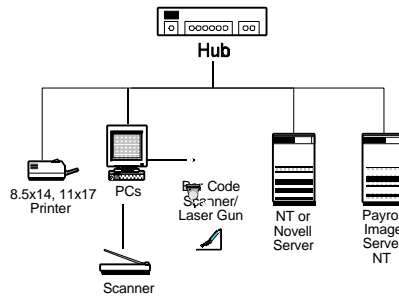
For any census tract in which the response rate is less than 100 percent after the initial response period, enumerators will perform nonresponse follow up (NRFU) (that is, contact the respondent and complete a census questionnaire) under the direct sampling strategy. In the direct sampling strategy, the Census Bureau will select a sample of nonresponding addresses in each census tract at the end of the initial response period. Enumerators will perform NRFU for each of the selected sample addresses. The addresses will be visited by an enumerator who will complete a questionnaire by personal interview.

**Local Census Offices (LCOs)** There will be approximately 475 LCOs which will be located within RCC areas and each LCO will report to a single RCC. LCOs will be opened exclusively to perform census field operations. The LCOs will provide the following support during census field operations:

- Produce enumerator maps and assignments;
- Conduct local recruiting;
- Conduct outreach and promotion;
- Conduct group quarters/special place/service-based enumeration operations;
- Coordinate military enumeration (if applicable);
- Conduct update/leave operation;
- Conduct list/enumerate operation;
- Conduct follow up enumeration;
- Manage field staff payroll and personnel administrative system.

### **LCO Architecture**

The following architecture will fully support LCO census field operations. Each LCO will consist of a 802.3 10BaseT Local Area Network (LAN) with a PC Windows NT or Novel server (1), PC-based desktop computers (13 PCs), and imaging workstations (2 PCs). Applications developed for the LCOs will be both batch oriented (e.g., offline printing, data/image capture) and client/server, accessing an Oracle RDBMS and PeopleSoft Payroll/Personnel software located at an RCC. Connectivity to the RCCs will be through a 64Kbps Frame Relay Link.



## LCO ARCHITECTURE

### 2.1.3 Decennial Field Operations

**Field Collection Control** The Decennial Field Interface (DFI) is the framework for **all** systems used in field collection control activities at the RCCs and CFOs for early census field operations, and for field collection control activities at the RCCs and LCOs for census field operations. The overall goal of this interface is to provide the necessary data in a timely fashion for the effective management of the field operations.

**The DFI consists of eight principal components:**

1. Operations Control System 2000 (OCS 2000);
2. Office Support Systems;
3. PAMS/ADAMS (covered in detail under Administrative Systems);
4. Geography Systems (covered in detail in Geography IT Operational Plan);
5. MIS (covered in detail under Administrative Systems);
6. Integrated Coverage Measurement (ICM) Control (covered in detail under Integrated Coverage Measurement)
7. Administrative Control Systems;
8. Staff Authorization.

All of the DFI components or sub-systems are Government developed using the following COTS software.

DFI Sub-system	COTS Software
OCS 2000	Unix, Oracle, Powerbuilder, C
Office Support Systems	Windows 95, CC Mail, Wordperfect, Excel
PAMS/ADAMS	Unix, Oracle, People Soft, COBOL
Geography Systems	GUSX and MIMprint
MIS	SAS
Integrated Coverage Measurement Control	--CASE Management, Unix, Oracle, Powerbuilder, C
Administrative Control Systems	Oracle, dBASE
Staff Authorization	Quattro Pro Spreadsheet

A key component of the DFI is the Operations Control System (OCS 2000) being developed by the Bureau expressly for field collection control activities. The OCS 2000 will capture and provide timely data necessary for the effective management of the RCCs, CFOs, and LCOs. Some of the specific functions of the OCS 2000 include the following:

- **Assignment tracking** - Maintains information on the progress of assignments;
- **Production printing/Reporting/Graphics** - Produces maps, listings, labels, and directories;
- **Group Quarters (GQ) Control** - Extracts Special Place/Group Quarters assignments, prints appropriate directories, listings, control records, updates the Master File, and controls the check in and checkout of work through the appropriate phases;
- **Housing Unit (HU) Control and Enumeration Operations** - Accesses files with listings, directories, and labels needed, tracks the universes and cases completed through check in and checkout procedure for such operations as Nonresponse Follow up and reinterview (a QA operation), Update/Enumerate, and List/Enumerate;
- **Be Counted** - Identifies targeted areas, distribution sites, and maintains an inventory of forms, questionnaires and promotional materials;
- **Management Activities** - Provides functions for managers that are not available to other staff. The specific functions to be included are being

identified, but currently include enabling operations, making changes to crew leader assignments, etc.

#### **2.1.4 Outreach Support System (OSS)**

The Address Listing operation is the first step in the development of the Census 2000 Outreach Support System (OSS). The OSS is designed to support presentations to local groups on the Decennial Census process and goals. This system will have the ability to present local profiles of census data so that presentations can be tailored to the local area. The OSS will reside in the RCC and will upload and download data to staff laptops as required. Laptops with 2 to 4 GB storage capacity and multimedia capacity are required.

### **2.2 Benefits**

It is anticipated that the data collection system for Census 2000 will result in the following benefits:

- Improve the timeliness and completeness of information available to the CFOs, LCOs, and Headquarters for managing the overall Census Operation;
- Reduce costs, by eliminating or scaling back on the hardware and software requirements for remote field offices (i.e., no automation in CFOs);
- Reduce technical support, by concentrating processing at the RCCs. This will also reduce software and hardware maintenance requirements;
- Improve the availability of data collected, by concentrating the storage and access to all field information at the RCCs.

### **2.3 Acquisition Strategies:**

**RCC Support for Early Census Field Operations and Census Field Operations** The BoC plans to obtain the necessary computer equipment and services for the RCCs through precompeted Government-Wide Agency Contracts (GWACs).

**LCO Support for Census Field Operations** In April 1998, we will begin developing the requirements for the full set of census operations to be conducted out of the LCOs, which will open in the Fall of 1999. A multi-functional team will be established to implement a streamlined acquisition process for the hardware and software needed in the LCOs.

## **3.0 Data Capture**

### **3.1 Background**

The data capture methodology implemented in Census 2000 will use the best state-of-the-art technology available. Furthermore, this technology will allow the use of more respondent-friendly questionnaires. Lastly, for the first time, both major components in the data capture area will be contracted out.

The two major components of data capture are:

1. **Data Capture System 2000 (DCS 2000)** The design and development of the data capture system;
2. **Data Capture Services Contract (DCSC)** The acquisition, build out, operation, and management of the Data Capture Centers (DCCs) that will house and support the operational areas of DCS 2000.

### **3.2 Data Capture System 2000 (DCS 2000)**

#### **3.2.1 Description**

Data capture begins with the check-in of incoming mail. Data Capture Centers (DCC) will receive respondent forms (i.e., mail return forms) and other types of forms including continuation, enumerator-filled, and group quarters forms from local census offices (LCOs), and Be Counted forms from the public.

The data capture process is the automatic or manual capture of respondent data from completed Census forms. Specifically, it involves the conversion of respondent answers and other control information on the census form to an electronic format suitable for computer processing. In DCS 2000, this will be accomplished by image scanning a form and then passing that image through optical mark recognition (OMR) and optical character recognition (OCR) engines. Via this process the image will then be converted into a computer-readable data format. When the OMR/OCR engines cannot successfully interpret the data on the form, a key from image (KFI) process will then be used to capture the data. All outputs from DCS 2000 are electronic and include regularly scheduled transmission of files containing check-in data, ASCII census data, and intermediate status data.

#### **Primary Goals of DCS 2000:**

- To identify respondents quickly so that followup activities can be conducted for non-respondents;
- To derive ASCII Title 13 data from the paper returns;
- To maintain intermediate status data, such as metrics on work in process and productivity rates.

#### **Physical Inputs consist of the following:**

- Returns in envelopes from the USPS;
- Returns from local census offices via overnight deliveries (LCOs).

### **Key Workload Assumptions for System Sizing:**

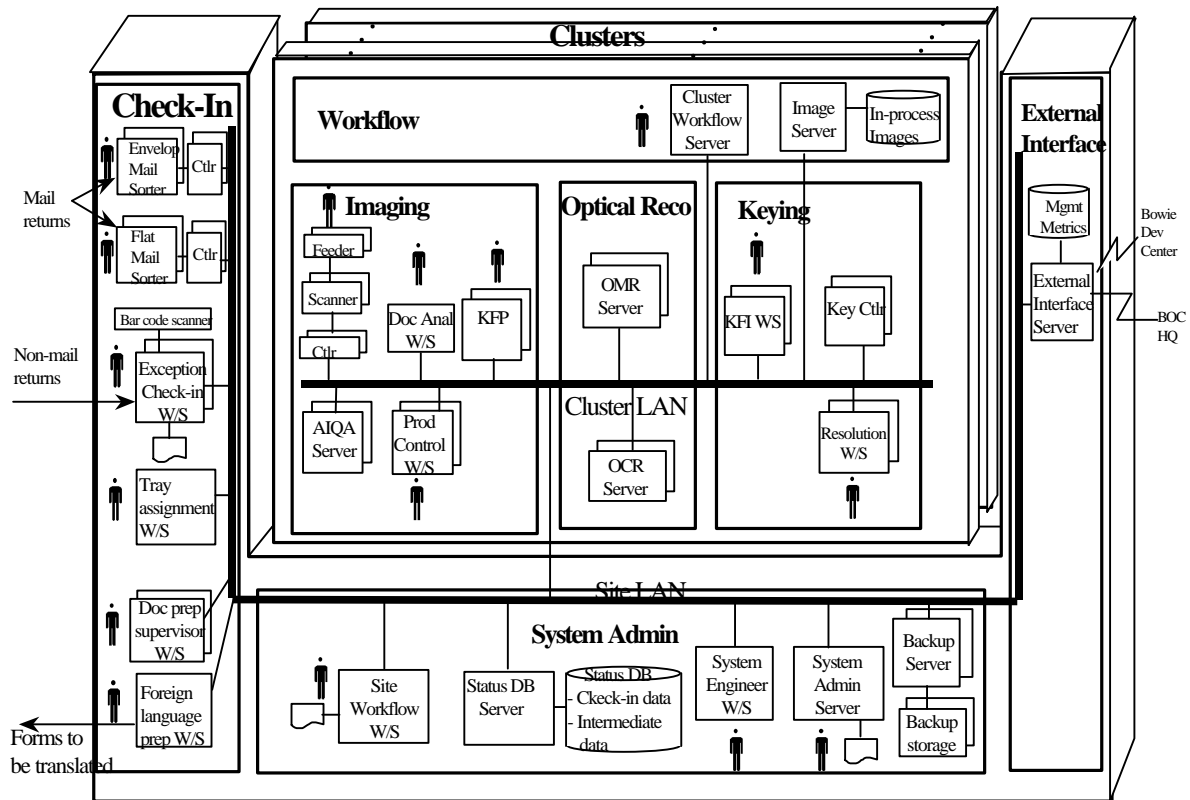
- The mail return rate will be 67%;
- A total of 125 million forms will be captured over 110 calendar days.

### **Architecture**

The system architecture is segmented at the highest level by separating functions driven by mail arrivals (which must be processed very quickly after receipt to provide check-in data to BoC) from those functions associated with the need to provide ASCII data from the content of the returns. The data for the latter functions can be processed in a more steady state environment, as long as priority processing deadlines are met. Functions relating to check-in processing, electronic interface with BoC headquarters and overall system administration are allocated to the site level for a DCC. The remaining functions, are those functions that extract data from paper forms and convert it to ASCII data, are handled by "clusters." Clusters are autonomous units of image processing constructed around the capacity of three scanners. Each DCC will have as many clusters as are necessary to process the assigned DCC workload.

The major DCS 2000 system components will run on Windows NT, using the COTS software detailed in the table below. Provided is an estimated range of the number of components that will be required in Census 2000. Final totals will not be determined until after the dress rehearsal.

## DCS 2000 System Architecture



### Components:

#### Site Level: DCS 2000 Subsystem Components:

- Check-In -> Sort incoming mail from USPS;
- Systems Administration -> Handle site workflow, systems administration

- and backup;
- External Interface -> link DCS 2000 with other BoC systems.

#### **Cluster Level” DCS 2000 Subsystem Components:**

- Imaging -> Scan documents and check for image quality;
- Optical recognition -> Read information from images / output ASCII data;
- Key from Image (KFI) -> Repair data that recognition engine could not interpret with a sufficient level of confidence;
- Audit Resolution -> Check for missing persons & data on the forms;
- Cluster Workflow -> Handle the workflow for the cluster.

#### **DCS 2000 COMPONENTS**

<b>DCS 2000 Component</b>	<b>National TOTAL (range)</b>	<b>COTS Software</b>
Sorters	40 - 50	None
Scanners	120 - 160	Scan Import, Form ID, ICE, Form QC, ICR3
Key Stations	1500-2200	Textware Form Client
Other Workstations	35- 55	Staffware Workflow Management
Recognition Servers	100-250	FAQSS (OMR), CGK, RecoStar, (OCR)
Other Servers	120 - 160	Staffware Workflow Mgt, Oracle
Backup Storage Automated Tape Libs	10	Platinum Net Archive

### **3.2.2 Benefits**

It is anticipated that DCS 2000, the data capture operation for Census 2000, will produce the following benefits:

- Ability to increase level of accuracy over time;
- Faster availability of data;
- Retention of images;
- Technology infusion;
- Potential for improved workflow and more flexible use of space and people.

### **3.2.3 Acquisition Strategy**

Using full and open competition, the BoC awarded this contract to Lockheed Martin in

### **3.2.4 Performance Measurement**

The following performance measurements were used at several demonstrations during the DCS 2000 development phase. During actual production, these measurements will be used to monitor system performance and to react to system problems. These measurements will also be used in the determination and validation of production system sizing. As a result, the DCS 2000 architecture and sizing will be adjusted based on these testing results.

#### **The Measurements:**

- Scanner capacity;
- Recognition servers capacity;
- Recognition accuracy/success rates (alpha and numeric reject rates);
- Keying rates;
- Frequency of forms that cannot be successfully scanned;
- Mail sorter capacity.

## **3.3 Data Capture Services Contract (DCSC)**

### **3.3.1 Description**

The Data Capture Services Contract (DCSC) refers to the infrastructure that will support the capture of data from forms for Census 2000. The DCSC will provide three vendor-owned data capture centers (DCCs), staff, office equipment and supplies, training, and procedures to process Census 2000 questionnaires using the technical solution provided by the DCS 2000 contractor. The fourth DCC will be located at the BoC's Jeffersonville, IN, site and will be managed by the Government.

#### **The DCSC Contract Provisions:**

- Obtain and build out the DCC space using BoC and DCS 2000 specifications;
- Equip the DCCs with all office equipment and supplies needed to operate a facility;
- Recruit, hire, train, and manage a large temporary workforce to support Census 2000 data capture operation;
- Operate DCS 2000 equipment;
- Provide facilities management.

#### **Operational Support Provided by DCCs:**

- Check in mail returns;
- Conduct data capture;
- Manage payroll and personnel administrative system.

### **3.3.2 Acquisition Strategy**

Using full and open competition, the BoC expects to award this contract in January of 1998.

### **3.3.3 Performance Measurement**

Most of the performance measures shown below are schedule and cost driven. Section C of the contract will provide the performance indicators and allowable variations. Upon contract award, processes to monitor overall performance will be established. This is in addition to specific performance measurements and indicators.

#### **The Measurements:**

- Facilities Preparation;
- Management and Operation Services;
- Data Capture Operations Services;
- Program Management
- All deliverables.

## 4.0 **DATA PROCESSING**

### 4.1 **Description**

The objective is to develop an effective and efficient automated data processing system for controlling, managing, and processing census 2000 data. It will include the necessary interactions with the Master Address File (MAF), Decennial Field Interface (DFI), Data Capture System 2000 (DCS 2000), and telephone questionnaire assistance (TQA), not only to control, accept, and store the data but also to provide the necessary computer processing to produce the census count for the President, and to produce Census data products. To achieve this objective there are a number of activities which must be completed. These activities have been broken down into the following three data processing or operational phases:

1. **Pre-census Activities.** Those activities required for converting the MAF into the decennial control data base that remains linked to both TIGER and the MAF. These activities will include form sampling (long or short), targeting identifications (for example, with foreign language questionnaires), and preparing the address files for printing on the questionnaires. Control information (both geographic and address related) will be provided for guiding both field canvassing and address capture processing through post-capture processing.
2. **Census Activities.** Activities concurrent with census data collection/capture will be those necessary to coordinate the check-in capture and storage of census questionnaire data from the multiple sources of collection (mailback, enumerator, Be counted questionnaires, and telephone), to define the responding/nonresponding universes, and to provide enumeration controls and workload to the field. Included in these activities will be the loading and updating of the central data bases for the storage of all census responses provided through the enumeration and data capture processes. This also includes the unduplication of questionnaires.
3. **Post-Census Activities.** Those activities necessary to prepare data from the original responses for release. These activities will include final unduplicating of multiple responses, editing and imputation, coding of write-in response data, estimation, tabulation recoding, and data disclosure avoidance.

In addition, the detailed data files will be prepared from information collected on the long-form questionnaires. These activities will include editing responses, applying statistical techniques to account for missing data, and applying weights to these sample records. The files will be provided to DADS for data dissemination.

The data processing operations will take place in Data Preparation Division (DPD) located in Jeffersonville, IN, and at Census Bureau Headquarters facilities. The data processing function of each of these infrastructure components is as follows:

Centralized data processing systems to manage and control all census activities and to process the data collected will be done at the Bowie Computer Center of the Census Headquarters and by the Data Processing Division (DPD) in Jeffersonville, In.

### **Headquarters Processing**

All data processing operations, other than address list capture and clerical coding, will be performed at Headquarters. These activities include the following:

- Decennial Management Controls (pre-census and census activities);
- Post Response Processing Systems (post-census activity).

### **DPD Processing**

The data processing operations will include:

- Address list capture operations (precensus activity);
- Clerical coding of questionnaires (post-census activity).

## **4.1.1 Headquarters Data Processing Applications**

### **4.1.1.1 Decennial Management Controls**

The Decennial Management Controls (DMC) is a complex network of operation controls that provides the necessary interactions with the Master Address File (MAF), the data collection sites, data capture centers, telephone questionnaire assistance centers, and multiple divisions in order to process the census.

Major pre-census system development activities:

1. **Decennial Master Address File (DMAF) Creation & Updates** The DMAF is a series of files that constitute the foundation of the Decennial master address listing control system that is used to control and track census operations;
2. **Decennial Field Interface (DFI)** The DFI is the framework for data collection control activities at the field offices. The DMC provides information in the form of files to the DFI to control data collection operations including update/leave, list/enumerate and nonresponse followup.
3. **Creation of Address File Tapes** Census forms will be printed, addressed and mailed by a contractor. Prior to mail out, files of addresses from the DMAF must be extracted and delivered to the contractor. Included in the files will be information necessary to place bar codes on the forms which indicate the geographic area, census housing unit ID, form type, and other information necessary to prioritize the data capture in the DCCs and to control the data collection activities.

### ***DMC major census system development activities:***

1. **Decennial Response File (DRF) Creation.** Every response to the census is stored in the DRF. The DRF is updated with data from the DCS 2000 and TQA. Results from the mail response data capture, Be Counted Forms, enumerator forms and group quarters forms will be stored to this file.
2. **Nonresponse Follow up (NRFU) Identification.** Based on the flags in the DMAF, the nonresponding housing units will be identified. The addresses for the nonresponding universe along with the surnames from the DRF will be sampled and fed to the Decennial Field Interface (DFI) for NRFU.

#### **4.1.1.1 Post Response Processing Systems**

The Post Response Processing Systems (PRPS) are a series of post data collection and capture operations. These include the operations required to resolve inconsistencies, code the write-in responses, edit the data, impute for missing data, apply estimations, recode for tabulation, apply disclosure avoidance, and preparations for loading into the data access and dissemination system.

PRPS major post-census applications:

1. **Automated Coding** is the process by which census codes will be assigned to responses written in on census forms. The automated portions of census data coding will occur at HQ. The Automated Coding process will begin with the development of the dictionary and reference files which will be stored in a database. Once the forms have been captured, the electronic files will be converted into ASCII data strings and sent to the DRF for post data capture processing. The appropriate ASCII data string will then be sent from the DRF to HQ's automated coding process where the data will be matched against dictionaries. Coding issues not resolved by the automated process will be sent to DPD (see description of Clerical Coding).
2. **General Coding** will involve assigning census codes for "other" race identification, as well as other household persons and relationships, and will occur at Headquarters. The General Coding data that are not assigned codes through the automated coding system will be clerically resolved by HQ's personnel. The General Coding application will utilize a client/server environment with a graphical user interface (GUI). General coding for the 100% data must be completed before the preparation of the apportionment counts due to the President in December 2000.

#### **4.1.1.3 Headquarters (HQ) Processing Architecture**

The following architecture will fully support HQ pre-census, census and post-census processing operations through 2001. All processing operations will be accomplished on the VAX Alphaserwer 4100 with 1 GB memory, and 750 GB raid storage. All of the data processing applications are Government developed using Digital Open VMS operating system, Fortran, C, C++, Borland Delphi, Turbo Pascal, and SAS languages. We estimate that the HQ programming staff will need to be supplemented with approximately forty contract programmer positions through 2000.

### **4.1.2 DPD Data Processing Applications**

#### **4.1.2.1 Address List Capture Operations**

Address List Capture Operations (ALCO), previously referred to as Pre-Census Address Listing Data Capture Operations, is a series of operations conducted in Jeffersonville. ALCO provides the data capture of the different address compilations and update/listing field operations. The data capture involves two distinct types of operations:

- **Address List Capture** is based on keying from the bound address registers or listing pages received from the data collection operation and local government officials, and includes the check-in, document preparation, keying, quality assurance and reporting subsystems. The workload for this operation is approximately 200,000 address registers (bound books) with over 23 million addresses. This is a key from paper operation which will use low-end servers configured with VT320 (ASCII) terminals for data entry. This operation will be done on an upgraded VAX based keying system currently being used in the Economic Census.
- **Address List Map Spot Digitizing**. Each address register includes a map that the data collection CFO personnel have used to place map spot indicators for each housing unit. Once DPD has checked in the address register, the map will be scanned on high-speed scanners, using technology provided by the DCS 2000 contractor. Using the resulting image, each map spot will be digitized and ASCII data will be sent to Geography Division for processing. The digitizing PC workstations (160) will be configured with TCP/IP software for network communications, and image viewing, cleanup, and workflow software. We estimate that approximately 600 temporary personnel will be hired to staff the ALCO keying and digitizing operation which will begin in early 1998.

#### **4.1.2.2 Clerical Coding**

As described earlier, automated coding will take place on HQ's systems, while any data records requiring clerical review will be forwarded to DPD for coding resolution. The major coding operations are described below.

1. **Geographic Coding** includes the Place of Birth (POB), the Place of Work (POW) and the Migration (MIG) coding subprocesses. The POW and POB ASCII string data will be sent to the HQ Geography Division. The data will then be matched against the appropriate dictionaries, and will also be matched against the TIGER database for geographical coding. Coding issues that are not resolved at the Geography Division will be sent to DPD. Those data records will be matched against the DPD's coding dictionary and against the TIGER database for clerical review and final coding resolution. Currently, this operation must be complete by 12/25/00. However, since these data are not needed for the apportionment counts, the BoC may decide to complete this coding later than December 2000. If this decision is made, then IT resources used in earlier operations (i.e., KFI, ICM) could be used for the Geographic Coding operation.
2. **Industry & Occupation (I&O) Coding** will involve coding the write in responses to selected questions on the census long form. I&O coding is a complex process that requires coding of multiple response items which are dependent on each other. The ASCII data string will be sent to the HQ for automated coding. The I&O Coding data that are not assigned codes at HQ, will be sent to DPD. These records will be matched against DPD's dictionaries for clerical review and coding resolution. These data are not needed for the apportionment counts and therefore, can be completed later than December 2000. IT resources used in earlier operations (i.e., KFI, ICM) may be used for the I&O Coding operation.

#### **4.1.2.3 DPD Processing Architecture**

The following architecture will fully support HQ pre-census, census and post-census processing operations through 2001. All processing operations will be accomplished on a VAX Alphaserver 4100 with 1 GB memory, and 250 GB raid storage. All of the data processing applications are Government developed using Digital Open VMS operating system, Fortran and C languages. The ALCO keying component will use VAX Alpha servers configured with VT320 (ASCII) terminals (600) using Viking software, Borland Delphi, Turbo Pascal, and C++. The ALCO map spot digitizing component will run on Windows NT on Pentium workstations (160). The ALCO map imaging component will use high-speed scanners (6) provided by the DCS 2000 contractor. The clerical coding operations will run on Windows NT on PCs (900).

## **4.2 Benefits**

Census 2000 data processing system benefits:

- The design configuration for Census 2000 takes advantage of readily available hardware and software, and meets the BoC's requirements without relying on a high degree of customization;
- Maximizes the use of commercial technology;
- Equipment can be reused as part of the corporate infrastructure after Census 2000;
- Faster availability of data.

## **4.3 Acquisition Strategy**

We have begun developing the requirements for the full set of data processing operations to be conducted in Headquarters and DPD. A multi-functional team has been established to implement a streamlined acquisition process for the hardware and software needed in HQ and DPD.

## **5.0 Data Access and Dissemination (DADS)**

### **5.1 Description**

The BOC chartered the development of DADS to facilitate the access and dissemination of official demographic and economic information. This interactive system uses the Internet technology as an access and delivery tool to allow both novices and experts efficient and cost-effective access to data generated by the BOC. DADS will provide the tabulation system to generate summary data, the publication system to provide summary data to out customers, and the Internet connection for the distribution of the tabulations and maps.

DADS will fully support the Decennial hardware and software requirements for production and electronic dissemination of Decennial data products. DADS IT plans provide for hardware and software support for delivery of:

- Internet and Intranet access;
- system functionality to build and deliver all of the 1998 Dress Rehearsal products, Census 2000 Standard Products, all Census 2000 PL 94-171 Products and associated geographic products;
- system functionality for users to create user-defined products via extraction and tabulation capabilities;and
- system functionality to produce/ press master CD-ROMS and print files for further production and distribution by other areas of the Bureau.

### **Architecture**

The current DADS architecture is based on a three-tier computing model. Each tier performs a specialized function and is connected to the other tiers via either the Internet or Census Intranet. The tiers are:

1. Client platforms
2. Application Server
3. Database Server and Map Server

There are three architecturally similar environments maintained. The Census Intranet environment supports internal Bureau of the Census users, referred to as the “Internal Production Environment” as well as the DADS development systems, referred to as the “Development Environment”. Census Intranet systems are protected from external access by the Census Internet firewall. The Internet environment supports external users outside of the Bureau of the Census and is referred to as the “External Production Environment”. The external systems all physically reside outside of the Census Internet firewall and do not allow access to internal Bureau of the Census systems.

### **Client Tier**

The client tier represents the systems and software used to access DADS by the user community. These systems are expected to vary greatly in their configurations, but will have the following characteristics in common:

1. Windows 3.1, Windows 95, Windows NT, UNIX, or any other operating system with a graphical user interface (GUI) capability
2. Support for a World Wide Web (WWW) browser that runs Java and JavaScript
3. At least 16MB of RAM
4. A connection to the Internet or Census Intranet

All DADS software will be implemented as Java Applets or HTML pages which will be run in the user's web browser. Applets and HTML pages will be downloaded from the DADS Application Server.

**Application Server.** The Application Server platform consists of hardware and software systems that implement the DADS application logic. The majority of the client platform's interaction will be with the DADS Application Server. The application server will support the display and interaction of the user interface, provide meta-data access, and will store various selections made by the users as they interact with DADS and construct queries. Processing on the application server will be characterized by a high volume of short transactions. Note that the application server platform will not execute queries against the DADS data sets such as Decennial, Economic, and the American Community Survey (ACS). This type of processing will be performed by the database server.

The following are the configurations currently in place for the Application Server:

Component	Development System	Internal Production System	External Production System
Model	SUN Ultra 170E	SUN Ultra 170E	SUN Ultra 170E
CPU's	1	1	1
RAM	512M	512M	512M
Disk Capacity	4G	22G	22G
Operating System	Sun Solaris 2.5.1	Sun Solaris 2.5.1	Sun Solaris 2.5.1

**Database Server.** Requests for access to the DADS datasets will be forwarded from the application server to the database server via high speed network (Fast Ethernet/100 Mbps). The database server will be exclusively dedicated to generation of result sets and transmission of the results to the clients.

The following are the configurations currently in place for the Database Server:

Component	Development System	Internal Production System	External Production System
Model	SUN Ultra 170E	SUN Enterprise 6000	SUN Enterprise 3000

CPU's	1	12	2
RAM	512MB	4G	1G
Disk Capacity	35G	420G	67G
Operating System	Sun Solaris 2.5.1	Sun Solaris 2.5.1	Sun Solaris 2.5.1

**Mapping Server** An additional component of the current DADS architecture is a mapping server. This system will provide processing dedicated to map based geography selection for queries as well as thematic map output. It is essential that we use a separate server for the mapping component so that we can tune each server to optimize performance for its specific application. The mapping platform will reside in the 3rd tier of the architecture (at the same level as the database server). Requests for mapping services will be initiated through the application server.

The following are the configurations currently in place for the Mapping Server:

Component	Development System	Internal Production System	External Production System
Model	SGI Indigo 2 High Impact	SGI Power Challenge XL	SGI Challenge DM
CPU's	1	12	4
RAM	256MB	4G	832MB
Disk Capacity	18G	320G	114G
Operating System	SGI IRIX (UNIX)	SGI IRIX (UNIX)	SGI IRIX (UNIX)

### **Database Architecture**

The database architecture is based on a distributed design. Database objects holding information such as meta-data tables, user query parameters, and other elements related to implementation of the user interface will all reside in an Oracle database on the application server. This database will be configured and tuned for On-Line Transaction Processing (OLTP) type of access, which will be characteristic of the type of transactions to be implemented in support of presenting the application user interface and construction of user queries. The Census datasets (ie. Decennial, Economic, ACS) will be housed in a database on the database server. This database will be configured and tuned for a large data warehouse environment, characterized by read only queries accessing large amounts of data.

### **Network Architecture**

As describe earlier in this document, The DADS system has a three tiered architecture. It is important that the communication between tiers two and three (Application Server, Mapping Server, and Database Server) be as fast as possible. Both the internal and

external systems (Tiers 2 and 3) are connected by dedicated fast Ethernet switches. The network protocols used to transfer data between tiers 2 and 3 are TCP/IP and Oracle's SQL\*Net. On the other hand, the network communication with the client (Tier 1) is difficult to control for the internal system and almost impossible for users outside the Census Bureau. Our system has been designed with this in mind and has therefore tried to make the client as "Thin" as possible (very little actual processing is done on the client). The client tier communicates with the DADS system using the HTTP and FTP network protocols.

### **Server Software**

The following application software components will be utilized in support of DADS on both the application and database server platforms:

- Oracle
- Netscape Enterprise Server
- Oracle Web Application Server V3.0
- JDBC software
- CORBA compliant Object Request Broker (ORB)

The mapping server will have the following software components:

- Oracle
- Spatial Database
- Oracle Web Application Server V3.0
- JDBC software
- CORBA compliant Object Request Broker (ORB)

### **Application Development Environment**

A number of software development tools will be utilized to facilitate the DADS application development process. The development environment is also using the Client (PC)/Server (UNIX Servers) model.

The Oracle Designer/2000 tool set will be used to support Computer Aided Software Engineering (CASE) activities during the requirements analysis and design phases of DADS. Designer/2000 components used include:

- Process Modeler
- Entity Relationship Diagrammer
- Function Hierarchy Diagrammer
- Designer/2000 Repository matrix tools
- Data Diagrammer

During the build phase, the following tools will be used:

- An Integrated Java Development Environment such as Symantec Visual Cafe or Microsoft Visual J++
- Oracle Procedure Builder

- Oracle SQL\*Plus
- Java Development Kit

## **5.2 Benefits**

By implementing a coordinated corporate approach to data access and dissemination throughout the BOC, DADS supports the BOC's strategic plan as follows:

- It will provide timely data while remaining innovative and highly sensitive to the needs of the customer,
- It will be designed to be customer-friendly and market-driven to provide products and services of superior value for our customers, and
- It will adopt the most effective and innovative processes to improve cost-effectiveness, cycle time, and quality performance in support of core business activities.

## **5.3 Performance Measurement**

DADS system performance measures are recorded in the DADS Test Plan and detailed test results from the 1996 DADS Prototype are available by request.

## **6.0 Integrated Coverage Measurement (ICM)**

### **6.1 Background**

After the completion of the decennial nonresponse followup (NRFU) operations, a follow-up survey for a representative sample of housing units across the nation will be conducted. This operation is referred to as the census quality check or ICM survey. It is designed to identify people missed within enumerated housing units and in entirely missed housing units, as well as people that were counted in the wrong place or those counted more than once during the initial census enumeration.

The objective of the Integrated Coverage Measurement (ICM) survey is to produce a “one number” census estimate of the U.S. population in Census 2000. This survey will improve accuracy, reduce costs, and eliminate the confusion and controversy caused by having more than a single “one number” measurement for the same population.

It is anticipated that the ICM survey will provide the means for including people of the types missed in earlier censuses and providing census results before legal deadlines. Finally, The BoC’s goals of ensuring a more complete count, having less differential among groups in the final numbers, and reducing costs are fully supported by the ICM implementation.

The ICM sample is compiled from a list of addresses created independently from the Master Address File (MAF). The final population estimates are the result of combining information collected from responses to the census, including mail returns and other opportunities to respond with results from the Census NRFU and ICM survey results.

#### **Three separate operational areas make up the ICM Survey.**

The areas of operation are:

- Data Collection;
- Data Capture;
- Data Processing.

### **6.2 ICM Data Collection**

#### **6.2.1 Description**

The ICM Data Collection is independent of all other data collection operations in Census 2000. Yet ICM will use the same infrastructure used in Census 2000 proper to complete its data collection. This infrastructure is fully described in the Data Collection section of this document.

The ICM Field Collection control component is part of Decennial Field Interface (DFI) data collection infrastructure. This field collection control component houses a series of Computer-Assisted Personal Interviewing (CAPI) control operations patterned after current survey operations. The control operations are:

- Sample Control
- Assignment Preparation/Tracking
- Work Unit Check-in
- Supervisory Review
- Management Reports
- Close-out

## **The ICM Data Collection operation contains two phases:**

### **Housing Unit Phase**

During this phase, housing units within sample blocks will be listed independently of the census and later matched to a census inventory of housing units. After reconciling the differences, the list of housing units confirmed to have existed within the sample blocks on Census Day will be compiled. This list will be used in conducting the ICM personal interviews.

### **Quality Check Person Interview**

In this phase, the interviewers will collect information about the current residents and anyone who has moved out of the sample blocks between Census Day and the time of the interviews. According to census residence rules, the interviewers (using Computer-Assisted Personal Interviewing devices) will then ask questions regarding alternate residences. This is to establish where people lived on Census Day.

## **6.2.2 Acquisition Strategy**

### **ICM Data Collection Equipment:**

In April 1998, we will begin developing the requirements for ICM interviewing using laptop computers. A multi-functional team will be established to implement a streamlined acquisition process for the hardware and software needed for this operation.

## **6.3 ICM Data Capture**

### **6.3.1 Description**

Each ICM enumerator will use a Computer-Assisted Personal Interview (CAPI) instrument, a laptop computer, in the data capture phase of the ICM operation. Via a modem, these data will be downloaded daily to the field office.

The manual systems used in previous censuses now automated through Computer-Assisted Personal Interviewing (CAPI) will provide the following benefits:

- Improve the timeliness and reliability of data;
- Improve the ability to monitor overall operation.

## **6.4 ICM Data Processing**

### **6.4.1 Description**

All of the ICM data processing operations will take place in either the Data Preparation Division (DPD) in Jeffersonville or at Census Headquarters in Bowie. The computer matching capability provided by the ICM System is designed to operational phase.

**The ICM Data Processing operation contains three phases:**

**Housing Unit Computer Matching**

An ICM independent listing (IL) of housing units is compared to a census listing for the same blocks for comparison. The result is a comprehensive list of housing units from both sources that will be used to drive the interviewing process.

**Matching Review & Coding System (MaRCS):**

**Housing Unit Clerical Matching**

In this phase, clerical matching will be performed in DPD using MaRCS software that supports automated ICM clerical data retrieval and data entry functions. The three steps needed to complete the Clerical Matching process are:

- Clerical Matching;
- Field follow up to resolve discrepancies;
- Final clerical matching.

**Matching Review & Coding System (MaRCS):**

**Person Matching**

The “person matching” process involves comparing the people counted in the ICM Survey to those enumerated in the actual census. After the computer matching, clerical matching will again be performed in DPD using the MaRCS software that supports automated ICM clerical data retrieval and data entry functions.

It is in this final step of the ICM Data Processing operation that the Census Bureau will produce estimates of people missed or duplicated in the census enumeration. These estimates will be used to update the critical census data files used in producing the final “one-number” Census 2000 counts.

## **6.4.2 Acquisition Strategy**

**ICM Data Processing Equipment:**

The requirements for ICM data processing will be included in the full set of data processing operations to be conducted in Headquarters and DPD. A multi-functional team has been established to implement a streamlined acquisition process for the hardware and software needed in HQ and DPD.

## **7.0 Telephone Questionnaire Assistance (TQA)**

### **7.1 Description**

The Telephone Questionnaire Assistance (TQA) operation will be contracted out to provide the public with convenient access to obtain general Census 2000 information, to obtain help in completing census forms, to request replacement forms, and to respond by telephone. The Census Bureau will implement an extensive TQA operation to support calls in English, Spanish, and other languages. We will have a well-publicized national, toll-free number and will use an automated touch-tone system that can handle a large number of calls concurrently.

The current plan is to acquire and implement a fully integrated nationwide network of call centers capable of supporting high-volume, limited duration inbound TQA calls from the general public as well as outbound follow-up calls to residential telephone numbers furnished by the Census Bureau.

The TQA acquisition will include, at a minimum, the following call-center services:

- Developing and implementing a system design;
- Combining computer-aided and live-operator services that include support in English, Spanish, and other languages to be determined;
- Routing calls, in a timely and efficient manner, across multiple networks if necessary;
- Responding with accurate and consistent information to all callers;
- Capturing the types of calls received;
- Capturing data from respondent interviews;
- Preparing data files and reports, and performing data storage and transmission;
- Reporting Management information system results on a real-time basis;
- Establishing a centralized referral system to resolve questions and communicate answers.

### **TQA Operations**

In the 1990 Census, temporary census Bureau staffs provided paper-based questionnaire assistance to the public by telephone and through walk-in assistance centers. Technological advances continue to broaden the ways for furnishing such help and for facilitating the reporting of data. For Census 2000, the BoC is committed to providing the simplest and most convenient mediums for questionnaire assistance and data reporting.

Based on the current state of the industry, we anticipate the selected solution will provide every operator with a networked computer with access to a full range of electronic reference materials and programs to conduct interviews electronically. This solution eliminates paper records and allows ASCII data to be transmitted directly to Headquarters for data processing.

## **TQA Architecture and Integration**

The TQA architecture will be provided by the contractor, however, the BoC will actually pay on a per call, per minute basis. This per call/per minute cost includes line charge, automated handling, labor and associated automation (i.e., a PC for each operator). We estimate that 9 to 11 million calls will be received at the call centers in Census 2000, with approximately 80 percent of those calls occurring between March 3 and April 8, 2000. The following presents our best estimate of the number of inbound calls, by type of call and average duration of each call:

Type of Call	Estimated Number of Calls (millions)	Average Duration of Call (minutes)
Informational	4.4 to 5.3	4
Forms Request	2.7 to 3.3	4
Respond by Telephone:		
Short Form	1.5 to 1.9	8
Long Form	0.4 to 0.45	35

We will require a flat ASCII format for data output. The contractor is responsible to maintain privacy and security in compliance with provisions of Title 13, United State Code, the Bureau of the Census Administrative Manual, and applicable sections of the Department of Commerce Handbook of Security Regulations and Procedures.

Based on information provided by the industry, we know that there is the likelihood that no single call-center will be able to undertake the expected workload. Therefore, we believe that the successful contractor will have to provide the following integration services:

- assembling and contracting with as many call-center firms as it will take to meet the requirement;
- developing and implementing a universal operational design for all subcontractors;
- designing and implementing a “seamless” telecommunications infrastructure;
- demonstrating, in advance of Census 2000 and by means of tests and simulations, that the call-center network fully meets BoC requirements;
- developing contingency plan and preparedness;
- managing and coordinating all preparations and functional operations of subcontractors;
- planning and implementing a training program for subcontracted call center staffs at all levels;
- developing and implementing a central mechanism for subcontractors to report problems and obtain resolutions;
- ensuring the provision of maintenance support services for all subcontracted call centers;
- establishing and operating a “command center,” for monitoring the operations and status of participating call centers;
- preparing, maintaining, and providing current Management Information System (MIS) reports.

## **7.2 Benefits**

The TQA:

- Allows easy access for the public to ask questions and to respond directly to the census via telephoning interviewing.
- Improves our ability to handle peak workload of telephone calls.
- Maximizes the use of commercial technology and services by utilizing existing long distance carrier and call center industry practices.
- Provides the ability to answer a much greater percentage of calls than in the 1990 Census. If respondents are able to get their questions answered easily and quickly, they are more likely to complete and mail in their questionnaire. This in turn may contribute to a higher response rate and reduces the cost of nonresponse followup.
- Telephone and electronic data summaries will provide more timely information on call volumes, calls completed, call delivery problems, staffing levels, etc.

## **7.3 Acquisition Strategy**

A multi-functional team has been established to implement a streamlined acquisition process for TQA services in Census 2000. The team is conducting market research. This information will assist the team in completing the TQA requirements definitions and analysis, and in developing the Statement of Need.

## **7.4 Performance Measurement**

The TQA specification requires the following service level: 90% of all incoming calls requesting live operator assistance will be answered by a live operator within 20 seconds.

## **ADMINISTRATIVE ACTIVITIES**

### **8.1 Background**

The Census Bureau has instituted several management initiatives (described below) to facilitate a more effective and efficient planning process for Census 2000 as well as its actual implementation.

### **8.2 Management Information System (MIS)**

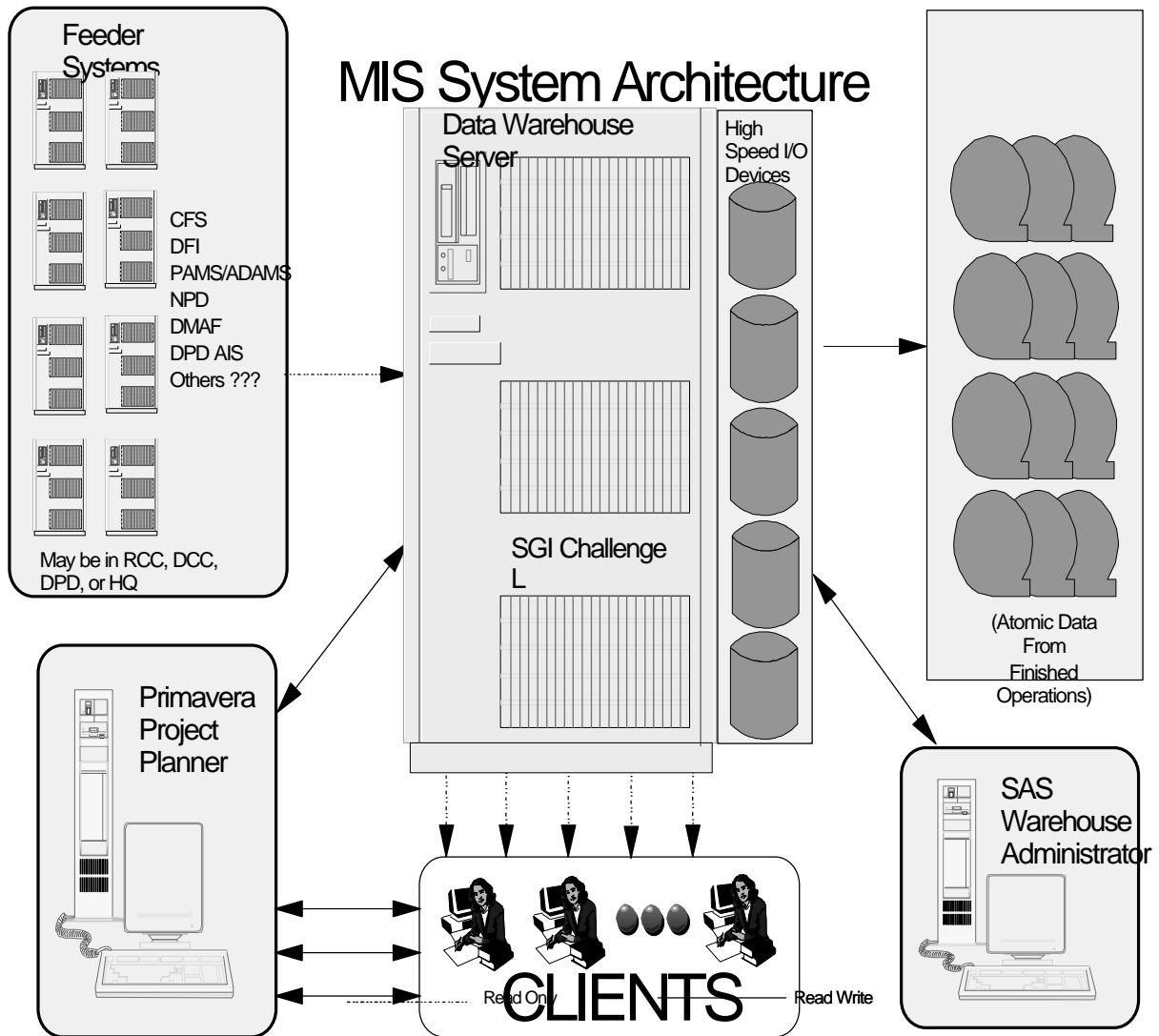
#### **8.2.1 Description**

2000 Census Management Information System (MIS) will be the official source of all information about the Census 2000, including schedule, performing divisions or organizations, budget, cost, and progress. In addition, the MIS will support cost modeling and other decisions support functions such as PERT, critical path analysis, what if analysis, and so forth. The MIS is a loosely coupled system consisting of Primavera Project Planner (P3) for the Master Activity Schedule (MAS) system and a cost and progress data warehouse and information delivery system which provides an enterprise-wide information system (EIS) to access the data warehouse. The data warehouse architecture will be a two-tiered client/server and will use the SAS Information Delivery System software products including: SAS/EIS, SAS/ASSIST, SAS/GRAPH, SAS/CONNECT, SAS/BASE, and others.

The data warehouse will reside at Headquarters. The client part of the cost and progress system for information delivery and data mining will reside on the users' desktops. The clients will access the data warehouse on a read-only basis and be able to analyze the information on the server or create data extracts to download for analysis. The clients also will have the P3 product installed on the desktop to read and update the schedule as appropriate. The schedule database will reside on an enterprise LAN file server.

The data warehouse will be updated daily with summary information from transactional systems located in the field (RCCs and DCCs) and Headquarters. Transactional systems which will feed cost and progress data to the warehouse include: the MAS, Decennial Field Interface (DFI), Pre-appointment Management System/Automated Decennial Administrative Management System (PAMS/ADAMS), Core Financial System (CFS), Decennial Master Address File (DMAF) with its updates, Address Listing Capture Operations (ALCO), and others to be defined. Summary level extracts from the feeder systems will be made nightly at a time which will have minimal impact on the transactional systems. As operations are completed, data will be moved to nearline storage.

Client (user) workstations will use the SAS software system to mine the data warehouse for operational cost and progress. The data architecture for the MIS is depicted below.



***MIS Architecture***

### **8.2.2 Benefits**

- Provide one official source of all information about Census 2000.
- Allow managers to analyze “near real time” information on all major Census 2000 operations.

### **8.2.3 Acquisition Strategy**

A multi-functional team has been established to implement a streamlined acquisition process for the hardware and software needed in HQ and DPD. Most of the MIS hardware and software has already been purchased using an approved Requirements Initiative; however, any future needs will be included in the streamlined process.

## **8.3 Pre-Appointment Management System/Automated Decennial Administrative Management System (PAMS/ADAMS)**

### **8.3.1 Description**

The DFI is the framework for **all** systems used in field collection control activities at the Regional Census Centers (RCCs), Census Field Offices (CFOs), and the Local Census Offices (LCOs) in support of the decennial census. The DFI consists of eight principal components; one of which is the Pre-Appointment System/Automated Decennial Administrative Management System (PAMS/ADAMS). This system will provide an automated enterprise-wide system which will support the hiring of employees, process personnel actions, pay employees, provide reports and data outputs, and maintain historical data. This system will support the personnel and payroll processing and information needs required by both management and operational personnel at all offices, including HQs, Regional Offices (ROs), RCCs, CFOs, and LCOs.

PAMS/ADAMS will support the following functions:

- Hiring of Employees. Early field operations in each RCC will require 2000 field employees for the urban operations and from 2000 to 2500 field employees in the rural CFO operations. To maintain this level of employment, there will need to be 30,000 to 40,000 applicants in the PAMS/ADAMS system in each RCC. The goal is to have a three day turnaround from the testing of an applicant to the job offer.
- Receipt/Capture/Processing Personnel Payroll Documents. The processing of personnel/payroll documents will take place in the RCCs. The processing will be done primarily on an image capture/Optical Character Recognition (OCR) subsystem which will be used to scan and to capture the payroll and personnel documents. The CFOs will ship applications and payroll forms to the RCCs via overnight delivery. Peak staffing will require two shifts of 20

administrative employees to centrally process the payroll for the CFOs.

PAMS/ADAMS has been sized to accommodate approximately 325,000 employees at peak employment levels in 12 regions working from as many as 500 LCOs. Overall, Census estimates that they will employ as many as 500,000 employees between January 1997 and October 2001. We estimate that there may be as many as eight applications accepted for every person hired (i.e., 4 million applications), and PAMS/ADAMS will accommodate this workload, as well as the more than 30 million time and expense forms which will be submitted.

### **PAMS/ADAMS Architecture**

Within each RCC, three high-end Unix servers will be responsible for storing and processing data for most of the DFI subsystems, one of these servers will support the PAMS/ADAMS component. The imaging subsystem includes three desktop scanners attached to three PCs in the RCCs, and one desktop scanner attached to a PC in the LCOs.

The PAMS/ADAMS components are using the following COTS software.

DFI Subsystem	COTS Software
PAMS/ADAMS	Unix, Oracle, People Soft, COBOL
Imaging component: <ul style="list-style-type: none"><li>• Scanner</li><li>• Key From Image (KFI)</li><li>• Storage/Retrieval of images</li></ul>	<ul style="list-style-type: none"><li>• DATACAP Taskmaster (includes imaging, OMR, and OCR)</li><li>• DATACAP PaperKeyBoard</li><li>• Watermark</li></ul>

### **8.3.2 Benefits**

- Decrease the time lag between the end of the payroll period and check disbursement.
- Improve the timeliness and completeness of information available to the CFOs, LCOs, and Headquarters for managing the overall Census Operation.
- Reduce costs, by eliminating or scaling back on the hardware and software requirements for remote field offices (i.e., no automation in CFOs).
- Reduce technical support, by concentrating processing at the RCCs. This will also reduce software and hardware maintenance requirements.
- Improve the availability of data collected, by concentrating the storage and access to all field information at the RCCs.

### **8.3.3 Acquisition Strategy**

### **Early Census Field Operations**

The BoC plans to obtain the necessary computer equipment and services for the RCCs through precompeted Government-Wide Agency Contracts (GWACs).

### **Census Field Operations**

The BoC conducted a competition between several vendors to provide the required hardware for the PAMS/ADAMS system. Based on a competitive process, we have contracted with Digital Equipment Corporation (DEC) for the Unix database servers to be used by PAMS/ADAMS in the RCCs. The PeopleSoft software was procured in 1996 during the development phase.

## **8.4 Decennial Applicant Name Check (DANC)**

### **8.4.1 Description**

The Decennial Applicant Name Check (DANC) is a related system, though not a component of PAMS/ADAMS. The following is the planned approach for the DANC workflow:

The PAMS will electronically send applicant data through the Decennial Applicant Name Check (DANC) at Census HQ. The data will be formatted and transmitted to the FBI in accordance with specifications provided by the FBI. The FBI will transmit name check results back to DANC. DANC will transmit result information back to the PAMS source in the RCC on a daily basis. This will allow Field staff to identify applicants who are eligible to be hired as quickly as possible. Applicants will be provided the opportunity to dispute unfavorable name check results by either providing fingerprints or disposition information. The DANC system is designed to cover all decennial applicants. All decennial applicants will be compared to the FBI index. It is expected that DANC must be able to process up to 100,000 applications per day during peak.

### **8.4.2 Benefits**

The DANC system is designed to ensure that all temporary Census Bureau employee applicants are properly screened. This includes census enumerator applicants whose names are retrieved from the local collection office (LCO) PAMS/ADAMS database and also decennial census support contract employees whose names may be received from sources other than the PAMS/ADAMS database. The purpose of the FBI name check process is to provide an adequate mechanism for ensuring that the 2000 Decennial Census is conducted by persons who meet Census Bureau requirements and to ensure that sensitive census data is handled by employees who meet Census Bureau requirements. Decennial Census enumerators will have direct contact with the public and census contract support employees will have access to sensitive census data.

### **8.4.3 Acquisition Strategy**

The DANC System is being developed in-house. It will operate on equipment used to support other Census 2000 systems.

## **8.5 Beta Site**

### **8.5.1 Description**

Although the Beta Site is not a separate system, it is a major component and is critical to the success of Census 2000. The Beta Site is an independent operation to test and assure the quality, completeness, and security of software systems, hardware systems, and network systems before they are released for production. The Beta Site is the central point for national decennial census problem resolution support for systems and operations conducted in the DCC, RCC, RO, LCO, DPD, and HQ. There are three major functions performed at the Beta Site:

- Beta Testing of Software and Systems;
- Integration of Components and Systems;
- National Support of Decentralized Systems.

#### **Beta Testing of Software and Systems**

The Beta Site will have a set hardware and software that is an accurate and current representation of the operating environment at the DCC, RCC, LCO, HQ, and DPD. The Beta Site:

- Provides software testing for security compliance;
- Provides a facility to test software releases on computer systems that are identical to production systems;
- Insures that new releases of software properly interface with related software applications;
- Provides for testing automated release, acceptance and installation for all software releases;
- Maintains complete version control of all deployed software (application and system);
- Provides real-time system monitoring.

#### **Integration of Components and Systems**

The Beta Site serves as the center for integration of the RCCs. The Integration Center provides the following services:

- Receives RCC equipment;
- Stages equipment;
- Installs operating systems, 3rd party software, and application software;
- Burns-in components;

- Stages for shipment, and prepare install tapes;
- Ships to the RCCs.

### **National Support of Decentralized Systems**

A National Support Center (NSC) will be established at the Beta Site. The functions of the NSC include the following:

- Provides a customer hot-line service in support of various decentralized operating units.
- Logs any problems that are reported and provides problem resolution.
- Controls the release of new software to various decentralized census operating units.
- Serves as a clearinghouse for hardware, software, clerical, personnel/payroll, and MIS problems.
- Provides real time performance monitoring of decentralized systems.

### **Beta Site Architecture**

The Beta Site hardware and software systems, that represent the operating environments at the DCC, RCC, LCO, HQ, and DPD, will be provided by the responsible functional area. Therefore, the architecture and the cost of components are fully represented in the other sections of this Plan. The Beta Site provides space and contractor support used to operate the Beta Test Center, the Integration Center, and the National Support Center. At peak operations in 2000, there will be approximately thirty contractor personnel covering seven days/twenty-four hour a day shifts.

## **8.5.2 Benefits**

- Reduce risk of systems failure.
- Provide a testing environment to assure that software systems work before being released into production in the field sites.
- Allow system developers and program managers to use the integration, testing, and support services to assure the successful operation of systems in the field.

## **8.5.3 Acquisition Strategy**

The BoC plans to obtain the necessary services for the Beta Site through precompeted Government-Wide Agency Contracts (GWACs).